

What is claimed is:

1. A method for speech processing in a distributed-speech recognition system having a
5 front-end and a back-end for recognizing words from speech signals, said method comprising the steps of:
extracting speech features from the speech signals, wherein the speech features contain a speech-to-noise ratio;
normalizing the speech features;
10 filtering the normalized speech features in a frequency domain; and
conveying the filtered speech features from the front-end to the back-end.
2. The method of claim 1, wherein the filtering step is carried out with a low-pass filter.
3. The method of claim 1, wherein the filtering step is carried out with a data-driven filter.
4. The method of claim 1, further comprising the step of converting the speech signals from
a time domain to a frequency domain prior to extracting the speech features.
- 20 5. The method of claim 4, further comprising the step of converting the speech signals to digital signals prior to converting the speech signals from the time domain to the frequency domain.
6. The method of claim 4, wherein the time-to-frequency domain conversion is carried out
25 by a Fast Fourier Transform in order to compute a magnitude spectrum and provide a plurality of magnitude spectrum values.

7. The method of claim 6, further comprising the step of non-linearly modifying the magnitude spectrum in order to generate a plurality of logarithmically-warped magnitude spectrum values.

8. The method of claim 7, further comprising the step of assembling the logarithmically-warped magnitude spectrum values in order to produce a set of feature parameters representative of the speech features.

9. A distributed speech recognition front-end comprising:

first means, responsive to a speech signal, for extracting speech features from said speech signal and for providing a first signal indicative of the extracted speech features;

second means, responsive to the first signal, for normalizing the extracted speech features and for providing a second signal indicative of the normalized speech features;

third means, responsive to the second signal, for filtering the normalized speech features in a frequency domain in order to reduce noise in the second signal and for providing a third signal indicative of the filtered speech features; and

means for conveying the third signal to a distributed speech recognition back-end in order for the back-end to recognize words representative of the speech signal from the third signal.

10. The front-end of claim 9, wherein the third means comprises a data-driven filter.

11. The front-end of claim 9, wherein the third means comprises a low-pass filter.

12. The front-end of claim 9, wherein the first means comprises:

a time-domain, pre-processing device to convert the speech signal to a digital signal;
a time-to-frequency domain conversion device to provide a set of magnitude spectrum values from the digital signal; and

an assembly device to assemble the set of magnitude spectrum values into the speech features.

13. The front-end of claim 9, wherein the third signal has a sampling rate, said front-end
5 further comprising means to reduce the sampling rate prior to conveying the third signal to the distributed signal recognition back-end.

14. A distributed speech recognition system for processing a speech signal, said system
comprising:

10 a front-end, responsive to the speech signal, for extracting speech features from the speech signal and for providing a first signal indicative of the extracted speech features; and

a back-end, responsive to the first signal, for recognizing words representative of the speech signals and for providing a second signal indicative of the recognized words, wherein the front-end has means to normalize the extracted-speech features and means to filter the
15 normalized speech features in order to reduce noise in the speech signal.

15. The system of claim 14, wherein the filtering means comprises a low-pass frequency filter.

20 16. The system of claim 14, wherein the filtering means comprises a data-driven filter.

17. A speech recognition feature extractor for extracting speech features from a speech signal, comprising:

25 a time-to-frequency domain transformer for generating spectral magnitude values in a frequency domain of the speech signal and for providing a first signal indicative of the spectral magnitude values;

a feature generator, responsive to the first signal, for generating a plurality of feature

vectors and for providing a second signal indicative of the generated speech features;

a normalizing means, responsive to the second signal, for normalizing the generated feature vectors and for providing a third signal indicative of the normalized feature vectors; and

a frequency filtering means, responsive to the first signal, for reducing noise in the normalized feature vectors and for providing the speech features indicative of the noise-reduction feature vectors.

18. The extractor of claim 17, wherein the frequency filtering means comprises a low-pass filter.

19. The extractor of claim 17, wherein the frequency filtering means comprises a data-driven filter.

20. A communication device having a voice input unit to allow a user to input speech signals to the device, and means for providing speech data to an external apparatus, wherein the external apparatus includes a distributed-speech recognition back-end capable of recognizing speech based on the speech data, said communication device comprising

a front-end unit, responsive to the speech signals, for extracting speech features from the speech signals for providing a first signal indicative of the extracted speech features, wherein the front-end includes:

means, responsive to the first signal, for normalizing the extracted-speech features for providing a second signal indicative of the normalized speech features, and

means, responsive to the second signal, for filtering the normalized speech features in order to reduce noise in the speech signals and for including the filtered speech features in the speech data..